RAW SEQUENCE LISTING
DATE: 01/14/2002
PATENT APPLICATION: US/09/972,912
TIME: 09:53:02

Input Set : **N:\Crf3\RULE60\09972912.raw**
Output Set: **N:\CRF3\01142002\I972912.raw**

SEQUENCE LISTING

```
   3 (1) GENERAL INFORMATION:
   5       (i) APPLICANT: SOPPET, DANIEL R.
   6                      RUBEN, STEVEN M.
   8      (ii) TITLE OF INVENTION: CARDIAC AND PANCREATIC PROTEIN AND GENE
  10     (iii) NUMBER OF SEQUENCES: 42
  12      (iv) CORRESPONDENCE ADDRESS:
  13          (A) ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
  14          (B) STREET: 1100 NEW YORK AVENUE, SUITE 600
  15          (C) CITY: WASHINGTON
  16          (D) STATE: DC
  17          (E) COUNTRY: US
  18          (F) ZIP: 20005-3934
  20       (v) COMPUTER READABLE FORM:
  21          (A) MEDIUM TYPE: Floppy disk
  22          (B) COMPUTER: IBM PC compatible
  23          (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  24          (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
  26      (vi) CURRENT APPLICATION DATA:
C--> 27          (A) APPLICATION NUMBER: US/09/972,912
C--> 28          (B) FILING DATE: 10-Oct-2001
  29          (C) CLASSIFICATION:
  31     (vii) PRIOR APPLICATION DATA:
  32          (A) APPLICATION NUMBER: 09/049,022
  33          (B) FILING DATE:
  35    (viii) ATTORNEY/AGENT INFORMATION:
  36          (A) NAME: STEFFE, ERIC K.
  37          (B) REGISTRATION NUMBER: 36,688
  38          (C) REFERENCE/DOCKET NUMBER: 1488.0620001
  40      (ix) TELECOMMUNICATION INFORMATION:
  41          (A) TELEPHONE: (202) 371-2600
  42          (B) TELEFAX: (202) 371-2540
  45 (2) INFORMATION FOR SEQ ID NO: 1:
  47       (i) SEQUENCE CHARACTERISTICS:
  48          (A) LENGTH: 2745 base pairs
  49          (B) TYPE: nucleic acid
  50          (C) STRANDEDNESS: double
  51          (D) TOPOLOGY: linear
  53      (ii) MOLECULE TYPE: DNA (genomic)
  56      (ix) FEATURE:
  57          (A) NAME/KEY: CDS
  58          (B) LOCATION: 233..1423
  60      (ix) FEATURE:
  61          (A) NAME/KEY: sig_peptide
  62          (B) LOCATION: 233..328
  64      (ix) FEATURE:
  65          (A) NAME/KEY: mat_peptide
```

Input Set : **N:\Crf3\RULE60\09972912.raw**
Output Set: **N:\CRF3\01142002\I972912.raw**

```
 66          (B) LOCATION: 329..1423
 69     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 71 GCAGCGGCAC GGCAGCAGCG GCAACAAGTG CCGGACTAGC AGAGCCAAGC CGGAGCAGTC     60
 73 CCTGCCGCCG ACACCGCCGG GCCGCCCGTC CGGGGCGCCG CGCATGGAGC GTGAGCTGCG    120
 75 GCGGTCGCCG GGGCTGAGCC GCGCGGAGCG CCGGGACGTG GATGTGGCCG CGATCTCCCG    180
 77 CCCTTGCCCC CGCCCCGCCG AGCTGGAGCT GCTCCCGGAC AAGATATGAG AA ATG         235
 78                                                            Met
 79                                                            -32
 81 AGT GTT GGA CGT CGA AGA ATA AAG TTG TTG GGT ATC CTG ATG ATG GCA      283
 82 Ser Val Gly Arg Arg Arg Ile Lys Leu Leu Gly Ile Leu Met Met Ala
W--> 83     -30                   -25                   -20
 85 AAT GTC TTC ATT TAT TTT ATT ATG GAA GTC TCC AAA AGC AGT AGC CAA      331
 86 Asn Val Phe Ile Tyr Phe Ile Met Glu Val Ser Lys Ser Ser Ser Gln
W--> 87 -15                     -10                   -5                1
 89 GAA AAA AAT GGA AAA GGG GAA GTA ATA ATA CCC AAA GAG AAG TTC TGG      379
 90 Glu Lys Asn Gly Lys Gly Glu Val Ile Ile Pro Lys Glu Lys Phe Trp
 91               5                   10                  15
 93 AAG ATA TCT ACC CCT CCC GAG GCA TAC TGG AAC CGA GAG CAA GAG AAG      427
 94 Lys Ile Ser Thr Pro Pro Glu Ala Tyr Trp Asn Arg Glu Gln Glu Lys
 95           20                  25                  30
 97 CTG AAC CGG CAG TAC AAC CCC ATC CTG AGC ATG CTG ACC AAC CAG ACG      475
 98 Leu Asn Arg Gln Tyr Asn Pro Ile Leu Ser Met Leu Thr Asn Gln Thr
 99           35                  40                  45
101 GGG GAG GCG GGC AGG CTC TCC AAT ATA AGC CAT CTG AAC TAC TGC GAA      523
102 Gly Glu Ala Gly Arg Leu Ser Asn Ile Ser His Leu Asn Tyr Cys Glu
103   50                  55                  60                  65
105 CCT GAC CTG AGG GTC ACG TCG GTG GTT ACG GGT TTT AAC AAC TTG CCG      571
106 Pro Asp Leu Arg Val Thr Ser Val Val Thr Gly Phe Asn Asn Leu Pro
107                   70                  75                  80
109 GAC AGA TTT AAA GAC TTT CTG CTG TAT TTG AGA TGC CGC AAT TAT TCA      619
110 Asp Arg Phe Lys Asp Phe Leu Leu Tyr Leu Arg Cys Arg Asn Tyr Ser
111               85                  90                  95
113 CTG CTT ATA GAT CAG CCG GAT AAG TGT GCA AAG AAA CCT TTC TTG TTG      667
114 Leu Leu Ile Asp Gln Pro Asp Lys Cys Ala Lys Lys Pro Phe Leu Leu
115           100                 105                 110
117 CTG GCG ATT AAG TCC CTC ACT CCA CAT TTT GCC AGA AGG CAA GCA ATC      715
118 Leu Ala Ile Lys Ser Leu Thr Pro His Phe Ala Arg Arg Gln Ala Ile
119           115                 120                 125
121 CGG GAA TCC TGG GGC CAA GAA AGC AAC GCA GGG AAC CAA ACG GTG GTG      763
122 Arg Glu Ser Trp Gly Gln Glu Ser Asn Ala Gly Asn Gln Thr Val Val
123 130                 135                 140                 145
125 CGA GTC TTC CTG CTG GGC CAG ACA CCC CCA GAG GAC AAC CAC CCC GAC      811
126 Arg Val Phe Leu Leu Gly Gln Thr Pro Pro Glu Asp Asn His Pro Asp
127                   150                 155                 160
129 CTT TCA GAT ATG CTG AAA TTT GAG AGT GAG AAG CAC CAA GAC ATT CTT      859
130 Leu Ser Asp Met Leu Lys Phe Glu Ser Glu Lys His Gln Asp Ile Leu
131               165                 170                 175
133 ATG TGG AAC TAC AGA GAC ACT TTC TTC AAC TTG TCT CTG AAG GAA GTG      907
134 Met Trp Asn Tyr Arg Asp Thr Phe Phe Asn Leu Ser Leu Lys Glu Val
```

```
135          180                      185                      190
137 CTG TTT CTC AGG TGG GTA AGT ACT TCC TGC CCA GAC ACT GAG TTT GTT        955
138 Leu Phe Leu Arg Trp Val Ser Thr Ser Cys Pro Asp Thr Glu Phe Val
139     195                      200                      205
141 TTC AAG GGC GAT GAC GAT GTT TTT GTG AAC ACC CAT CAC ATC CTG AAT       1003
142 Phe Lys Gly Asp Asp Asp Val Phe Val Asn Thr His His Ile Leu Asn
143 210                      215                      220              225
145 TAC TTG AAT AGT TTA TCC AAG ACC AAA GCC AAA GAT CTC TTC ATA GGT       1051
146 Tyr Leu Asn Ser Leu Ser Lys Thr Lys Ala Lys Asp Leu Phe Ile Gly
147                      230                      235              240
149 GAT GTG ATC CAC AAT GCT GGA CCT CAT CGG GAT AAG AAG CTG AAG TAC       1099
150 Asp Val Ile His Asn Ala Gly Pro His Arg Asp Lys Lys Leu Lys Tyr
151                      245                      250              255
153 TAC ATC CCA GAA GTT GTT TAC TCT GGC CTC TAC CCA CCC TAT GCA GGG       1147
154 Tyr Ile Pro Glu Val Val Tyr Ser Gly Leu Tyr Pro Pro Tyr Ala Gly
155          260                      265                      270
157 GGA GGG GGG TTC CTC TAC TCC GGC CAC CTG GCC CTG AGG CTG TAC CAT       1195
158 Gly Gly Gly Phe Leu Tyr Ser Gly His Leu Ala Leu Arg Leu Tyr His
159     275                      280                      285
161 ATC ACT GAC CAG GTC CAT CTC TAC CCC ATT GAT GAC GTT TAT ACT GGA       1243
162 Ile Thr Asp Gln Val His Leu Tyr Pro Ile Asp Asp Val Tyr Thr Gly
163 290                      295                      300              305
165 ATG TGC CTT CAG AAA CTC GGC CTC GTT CCA GAG AAA CAC AAA GGC TTC       1291
166 Met Cys Leu Gln Lys Leu Gly Leu Val Pro Glu Lys His Lys Gly Phe
167                      310                      315              320
169 AGG ACA TTT GAT ATC GAG GAG AAA AAC AAA AAT AAC ATC TGC TCC TAT       1339
170 Arg Thr Phe Asp Ile Glu Glu Lys Asn Lys Asn Asn Ile Cys Ser Tyr
171                      325                      330              335
173 GTA GAT CTG ATG TTA GTA CAT AGT AGA AAA CCT CAA GAG ATG ATT GAT       1387
174 Val Asp Leu Met Leu Val His Ser Arg Lys Pro Gln Glu Met Ile Asp
175          340                      345                      350
177 ATT TGG TCT CAG TTG CAG AGT GCT CAT TTA AAA TGC TAAAATAGAT           1433
178 Ile Trp Ser Gln Leu Gln Ser Ala His Leu Lys Cys
179     355                      360                      365
181 ACAAACTCAA TTTTGCATAG AAAGGTGTAT TTTGAATAGT TCCCATGTTG TGTTCTCACA     1493
183 TTAGAGTAAT TTCTATATTA AACCATGAAA ATTGCCTTTA TGAGTGATAC CCATTTGAGG     1553
185 GCCTCTAAAC CCTTCAATTT GGTACTCACG TGAAGAGGGA AAGCGGAAGA TGGTAATTTT     1613
187 TTTTTATGGA TGATATGGCA GGATGATTGG TTCTGATCTT ACCGGCTAGT GGTCATTTTT     1673
189 AAAAAACTTG TACCCTCTTA TCTGAAATCC TGTTTCTGGA ATTTGGCCAT TTTAAGTGAT     1733
191 TTTGTTTGCC CTCTTCTATA ATATTCCTAC TTCCCATAAT AATGACTGAT TTATTTGTAA     1793
193 TTCAGGTATT TATAAACCTA TTGGCTACAA AGACTTTGTT AAACATTATC CAGTGGTTTT     1853
195 CGTGAAATGG AATTATGTTT ATTTTTATGG GATTTGGGTA AATTTTAAAT TGTCTAGAAA     1913
197 ACTGAAATTT CAGTTGTCAG TTGTGGAATT CAGTTTTTCA ATTGTGGAAA TTTCCTGCCA     1973
199 CCCCAACAGT ATTTTTGTGT GTTAATTAAT TTTGCAAAAT GAGAATCATG GTGTGACACT     2033
201 CATCTAATTT ATCTTGTTGT GATGTTATGG TCATAATAAG GAGAAAGAGG GTTTAATTTT     2093
203 TCTTGTATTT GGTTTCCTGG TGGTATCATA GTGTAATTTT AGTATTTGAA AATCAGTGTG     2153
205 ATTCCTTAAT GGCCAACTGA AGATTGAATT GCCGCTAACA ACCATATCGT GTTAGTGAAT     2213
207 TTTCAATATG GACCAGGAAG GCATATGTAT TTTGAACTTG AGTGAAAAGG TTGAAGTTAC     2273
209 AGACTTTTGC ATAGATGGTT TGTCAATTTA AAATTCCAGA ATTATTATT GCCATATTTT     2333
```

Input Set : **N:\Crf3\RULE60\09972912.raw**
Output Set: **N:\CRF3\01142002\I972912.raw**

```
211 CACATGCTGC TTATACAAGA TTATTATTGA GTAGTAACTG TTCCCTGTCT ATGTAGAAGT   2393
213 GCCTGTGTTT TTATTTATTG TTCCAGATCA AAGACCAAAA CATTTTCTTA AATATCTCTT   2453
215 ATGTAATATT TTATTTGTAT ACAGTGTTGT TGATGAAATA TTTAACTAGA GCATGATATT   2513
217 TTAAATGTTA AGGTGTAACA TATGTTAAAT AAAACTGTTA TTTTTGAATT TTAAAATTTG   2573
219 TTTTTTGGGG GTATGAACTA CTAGAGTTTA AAATTCTGCC AAACTATTAC TTATATGTAC   2633
221 TATTGTGTAA CATACTTTCT TGAAATATTT TTGTTTATAG AATTGAAGGT TCTTATCAGA   2693
223 TGGGATACTG GGGATTATAA ACAATGGAAA TAAAGCCACT GTATTTTAA AA           2745
226 (2) INFORMATION FOR SEQ ID NO: 2:
228       (i) SEQUENCE CHARACTERISTICS:
229             (A) LENGTH: 397 amino acids
230             (B) TYPE: amino acid
231             (D) TOPOLOGY: linear
233       (ii) MOLECULE TYPE: protein
235       (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
237 Met Ser Val Gly Arg Arg Arg Ile Lys Leu Leu Gly Ile Leu Met Met
238 -32      -30              -25              -20
240 Ala Asn Val Phe Ile Tyr Phe Ile Met Glu Val Ser Lys Ser Ser Ser
241      -15              -10              -5
243 Gln Glu Lys Asn Gly Lys Gly Glu Val Ile Ile Pro Lys Glu Lys Phe
244   1              5              10              15
246 Trp Lys Ile Ser Thr Pro Pro Glu Ala Tyr Trp Asn Arg Glu Gln Glu
247              20              25              30
249 Lys Leu Asn Arg Gln Tyr Asn Pro Ile Leu Ser Met Leu Thr Asn Gln
250              35              40              45
252 Thr Gly Glu Ala Gly Arg Leu Ser Asn Ile Ser His Leu Asn Tyr Cys
253      50              55              60
255 Glu Pro Asp Leu Arg Val Thr Ser Val Val Thr Gly Phe Asn Asn Leu
256   65              70              75              80
258 Pro Asp Arg Phe Lys Asp Phe Leu Leu Tyr Leu Arg Cys Arg Asn Tyr
259              85              90              95
261 Ser Leu Leu Ile Asp Gln Pro Asp Lys Cys Ala Lys Lys Pro Phe Leu
262              100             105             110
264 Leu Leu Ala Ile Lys Ser Leu Thr Pro His Phe Ala Arg Arg Gln Ala
265              115             120             125
267 Ile Arg Glu Ser Trp Gly Gln Glu Ser Asn Ala Gly Asn Gln Thr Val
268      130             135             140
270 Val Arg Val Phe Leu Leu Gly Gln Thr Pro Pro Glu Asp Asn His Pro
271 145             150             155             160
273 Asp Leu Ser Asp Met Leu Lys Phe Glu Ser Glu Lys His Gln Asp Ile
274              165             170             175
276 Leu Met Trp Asn Tyr Arg Asp Thr Phe Phe Asn Leu Ser Leu Lys Glu
277              180             185             190
279 Val Leu Phe Leu Arg Trp Val Ser Thr Ser Cys Pro Asp Thr Glu Phe
280              195             200             205
282 Val Phe Lys Gly Asp Asp Asp Val Phe Val Asn Thr His His Ile Leu
283      210             215             220
285 Asn Tyr Leu Asn Ser Leu Ser Lys Thr Lys Ala Lys Asp Leu Phe Ile
286 225             230             235             240
288 Gly Asp Val Ile His Asn Ala Gly Pro His Arg Asp Lys Lys Leu Lys
```

Input Set : **N:\Crf3\RULE60\09972912.raw**
Output Set: **N:\CRF3\01142002\I972912.raw**

```
289                    245              250              255
291 Tyr Tyr Ile Pro Glu Val Val Tyr Ser Gly Leu Tyr Pro Pro Tyr Ala
292                260              265              270
294 Gly Gly Gly Gly Phe Leu Tyr Ser Gly His Leu Ala Leu Arg Leu Tyr
295            275              280              285
297 His Ile Thr Asp Gln Val His Leu Tyr Pro Ile Asp Asp Val Tyr Thr
298     290              295              300
300 Gly Met Cys Leu Gln Lys Leu Gly Leu Val Pro Glu Lys His Lys Gly
301 305            310              315                320
303 Phe Arg Thr Phe Asp Ile Glu Glu Lys Asn Lys Asn Asn Ile Cys Ser
304            325              330              335
306 Tyr Val Asp Leu Met Leu Val His Ser Arg Lys Pro Gln Glu Met Ile
307            340              345              350
309 Asp Ile Trp Ser Gln Leu Gln Ser Ala His Leu Lys Cys
310         355              360              365
```

312 (2) INFORMATION FOR SEQ ID NO: 3:

314   (i) SEQUENCE CHARACTERISTICS:

315   (A) LENGTH: 323 amino acids

316   (B) TYPE: amino acid

317   (C) STRANDEDNESS: Not Relevant

318   (D) TOPOLOGY: linear

320 (ii) MOLECULE TYPE: protein

325 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```
327 Gln Ser Lys His Arg Lys Leu Leu Leu Arg Cys Leu Leu Val Leu Pro
328 1               5               10              15
330 Leu Ile Leu Leu Val Asp Tyr Cys Gly Leu Leu Thr His Leu His Glu
331            20              25              30
333 Leu Asn Phe Glu Arg His Phe His Tyr Pro Leu Asn Asp Asp Thr Gly
334            35              40              45
336 Ser Gly Ser Ala Ser Ser Gly Leu Asp Lys Phe Ala Tyr Leu Arg Val
337            50              55              60
339 Pro Ser Phe Thr Ala Glu Val Pro Val Asp Gln Pro Ala Arg Leu Thr
340 65              70              75              80
342 Met Leu Ile Lys Ser Ala Val Gly Asn Ser Arg Arg Arg Glu Ala Ile
343            85              90              95
345 Arg Arg Thr Trp Gly Tyr Glu Gly Arg Phe Ser Asp Val His Leu Arg
346            100             105             110
348 Arg Val Phe Leu Leu Gly Thr Ala Glu Asp Ser Glu Lys Asp Val Ala
349            115             120             125
351 Trp Glu Ser Arg Glu His Gly Asp Ile Leu Gln Ala Asp Phe Thr Asp
352     130             135             140
354 Ala Tyr Phe Asn Asn Thr Leu Lys Thr Met Leu Gly Met Arg Trp Ala
355 145             150             155             160
357 Ser Glu Gln Phe Asn Arg Ser Glu Phe Tyr Leu Phe Val Asp Asp Asp
358            165             170             175
360 Tyr Tyr Val Ser Ala Lys Asn Val Leu Lys Phe Leu Gly Arg Gly Arg
361            180             185             190
363 Gln Ser His Gln Pro Glu Leu Leu Phe Ala Gly His Val Phe Gln Thr
364            195             200             205
```

L:27 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:}
L:28 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:83 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:87 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1